Report to the Board of Trustees of the California State Bar
Final Report on the Standard Setting Study and Public Comments Regarding Pass Line Options

September 5, 2017

I. Executive Summary

In February 2017, the Supreme Court of California (Court) called for the State Bar of California (Bar) to undertake a "thorough and expedited study" of the pass rate on the California State Bar Exam (CBX). The Court directed the Bar to ensure that the study includes:

> "(1) identification and exploration of all issues affecting California bar pass rates; (2) a meaningful analysis of the current pass rate and information sufficient to determine whether protection of potential clients and the public is served by maintaining the current cut score; and (3) participation of experts and stakeholders in the process, including psychometricians, law student representatives and law school faculty or deans."[1]

To accomplish this, the Bar organized the work into four interconnected studies – a Standard Setting Study, Content Validation Study, and two studies focused on student performance on the bar exam – and contracted with nationally recognized experts in the fields of professional certification, testing, and psychometric evaluation. For the Standard Setting Study, Chad Buckendahl, Ph.D., conducted the research with additional consultants contracted to observe and review his work. Those reviewers included Mary Pitoniak, Ph.D., a nationally recognized expert in standard setting, and Tracy Montez, Ph.D., Chief of Programs and Policy Review at the California Department of Consumer Affairs, which includes the Office of Professional Examination Services.

In addition to the external consultation provided by these subject-matter experts, the Standard Setting Study was conducted with extensive stakeholder engagement including: bi-weekly conference calls with deans of California's law schools (ABA and California accredited and non-accredited); and bi-weekly conference calls with a working group comprised of representatives of the Admissions and Education (A&E) Committee of the Bar's Board of Trustees, the Committee of Bar Examiners (CBE) and Supreme Court staff.

Dr. Buckendahl's report, *Conducting a Standard Setting Study for the California Bar Exam* (Standard Setting Study) was submitted to the A&E Committee and the CBE on July 31, 2017.[2] That report was accompanied by a staff report recommending the release of the Standard

[1] Attached as Appendix A.
[2] Attached as Appendix B.

Setting Study for public comment and presenting two options: maintaining the current bar exam cut score of 1440,[3] or reducing the cut score to 1414 on an interim basis to be used exclusively for the July 2017 administration of the CBX. The A&E Committee and the CBE accepted the report and released it for public comment immediately following the meeting.

During August the Bar solicited additional input from stakeholders, holding two full-day sessions – one day each in northern and southern California – to receive public comments, and distributing an online survey to licensed attorneys and applicants for admission. In response to these calls for public input, the Bar received:

- Over 5,000 public comments on the Standard Setting Study;
- Over 34,000 survey responses from licensed attorneys in California, approximately 15 percent of the population of active and inactive attorneys to whom the survey was distributed; and
- Over 4,000 survey responses from applicants who took the July 2017 bar exam, approximately 45 percent of the applicants to whom the survey was distributed.

A report on these comments was presented to the Law School Council, an advisory body to the CBE, on August 30, and to the A&E Committee and CBE at a joint meeting on August 31. In addition to the option of lowering the cut score to 1414, a third option of 1390 was presented for consideration for the A&E Committee and CBE joint meeting. After reviewing and discussing the report, the Law School Council voted to recommend that the Supreme Court consider adoption of an interim CBX cut score within the range of 1350 to 1390; the CBE voted to recommend maintaining the pass line at 1440 until more information can be collected and evaluated.

These extraordinary measures have been undertaken to ensure that the process of evaluating the bar exam pass line is analytically rigorous, inclusive, and transparent. In addition to summarizing and responding to survey results and public comments, this report provides: deep background on the history and evolution of the CBX; a detailed overview of the Standard Setting Study; simulations of the impact of different cut scores on pass rates as applied to 2008 and 2016 CBX results; and a discussion of a number of policy issues related to the impact of the cut score. This report concludes by recommending that the Board consider referring a third option to the Court regarding the cut score on the bar exam: a total scaled score of 1390 which remains within the confidence interval of two standard errors below the median score determined through the standard setting exercise. Thus, the options presented for consideration following the period of public comment are:

[3] Unless otherwise noted, this report will refer to total scaled scores when discussing the cut score. Throughout the report the terms “cut score” and “pass line” are used interchangeably.

1. No change to the current CBX cut score; or
2. Adopt an interim cut score of 1414; or
3. Adopt an interim cut score of 1390.

II. A Brief History of the Bar Exam in California

The legal profession in the United States is regulated by state governments leading to a broad diversity of rules and practice governing admission to the bar. In a majority of states only students who have studied at schools approved by the American Bar Association (ABA) may take the exam. Other states allow students who have earned degrees from state accredited, unaccredited, and correspondence courses to take the bar exam. In a small number of states it is still permissible to take the bar exam after "reading the law" – serving a form of apprenticeship under a judge or practicing attorney – a practice that developed in common law countries prior to the advent of law schools.[4]

In addition to setting the requirements for legal education of those who may take the bar exam, all states require that candidates meet minimum standards of character and fitness – known as a "moral character" determination in California – and almost all states require that applicants to the bar pass the Multistate Professional Responsibility Examination. Thus, while passing the bar exam is an essential prerequisite to becoming an attorney in the United States, it is not the only prerequisite.

California is one of the least restrictive states in setting the legal education requirements for who may take the bar exam. In California students who attend ABA accredited, state accredited, or unaccredited law schools may sit for the CBX. Students who studied through correspondence courses may also take the exam. California is also one of only five states that allows applicants to take the bar exam after reading the law.

Historical Evolution of the California Bar Exam

During the last century the CBX has undergone numerous changes. The number, type, and content of questions, have all changed. The grading process, amount of time allotted to applicants to answer questions, the weighting of different portions of the exam, and rules related to grading and reviewing exams that narrowly miss the pass line, have all also been modified. The one thing that has been relatively constant throughout this period, however, is the pass line which has remained at, or very close to, 70 percent.

[4] Wikipedia, "Admission to the Bar in the United States," and "Reading Law."

The first bar examinations in California were administered in 1920 by a Board of Bar Examiners composed of members appointed by the Supreme Court. The Board had been established in 1919, eight years before the State Bar was created. When the State Bar came into existence in 1927, the Committee of Bar Examiners was established to develop and implement the admission requirements.[5]

From 1932 until 1972 the CBX consisted of about 20 essay questions and applicants needed an average score of 70 percent of the highest possible grade to pass.[6] In 1972, the Multi-State Bar Exam (MBE) – a 200 item, multiple-choice test developed by the National Conference of Bar Examiners – was added to the CBX and given a weight of 30 percent of the total value of the exam with the essay portion of the exam at 70 percent. The pass line for the combined essay and MBE remained at 70 percent.

Though it is not well documented, at some point prior to the adoption of the MBE, the number of essay questions was reduced from 20 to 12. The twelve essay questions were drawn from three sections of five questions each and applicants were required to answer four of the five questions in each section and allotted just over 52 minutes per question to complete their answers.

In 1978 a number of additional changes were introduced to the CBX. The number of essay questions was reduced again, this time from 12 to nine, but applicants were required to answer all of the questions. The amount of time allotted to answering the questions was increased to allow a full hour per question. The weighting of the MBE and essay questions was also shifted so that the MBE counted for 40 percent of the total score and the essays counted for 60 percent. In addition, during this period a number of modifications were made to the grading process, in part to save time on the grading,[7] but also to provide a "second look" for exams that fell within the range of 67.3 percent to 70 percent.

In 1981 a "bifurcation" rule was implemented which allowed applicants to repeat only one part of the exam if they failed overall but passed one portion. Applicants were able to bifurcate their results as late as the July, 1985 exam, but this practice was discontinued because it was found actually to *decrease* rather than increase passing rates.[8] In 1984 the practice of sampling three essay exams was also discontinued because it was determined that the labor savings of reading

[5] Report of California Survey Committee," published by The State Bar of California, 1933.

[6] The historical development and evolution of the bar exam in California are not well documented. This section relies heavily on "History of General Bar Examination Structure and Pass/Fail Rules," Stephen Klein, Ph.D., July 9, 2011

[7] Exams were graded initially by randomly selecting the responses from one essay test session – three questions – grading these and then evaluating the applicants score in conjunction with the applicant's score on the MBE. If that score "was high enough to virtually assure they would pass if all their essay answers were read," then the applicant was admitted to the bar without grading the remaining essays (Klein, 2011, p. 2) If, however, the randomly selected essays did not appear to guarantee a passing score, then all of the remaining questions were graded to determine if the applicant passed.

[8] Report on the impact of bifurcation of exam results attached as Appendix C.

only a sample of a small proportion of the total exams had a negligible impact on the cost of administering the exam.
In 1983 a Performance Test section was added to the CBX. While this initially included a set of multiple-choice questions and scoring separate from both the MBE and essay portions of the exam, by 1985 the multiple choice portion of the Performance Test was eliminated and by 1987 the essay and Performance Test questions were combined into a single, written score.

In February 1985, the CBE recommended that the essay and performance test portions of the CBX be scaled to the MBE. Scaling the written portions of the examination to the MBE was a procedure that was used by many other states and recommended by the National Conference of Bar Examiners. The goal of scaling, in part, is to ensure that the level of difficulty of the exam remains relatively constant from one administration to another.

The CBE's psychometric consultant at the time, Stephen P. Klein, Ph.D., also prepared several reports and made multiple presentations on the subject of scaling the essay and performance test scores to the MBE, which included both pros and cons for adopting the new scaling method.

The CBE adopted in principle scaling the essay and performance test portions to the MBE during its July 1985 meeting. However, because the scoring was changing, a new pass line had to be determined. The CBE took the scaling recommendation to the State Bar's Board of Trustees and at its April 1986 meeting, the Board of Trustees approved the request to begin scaling the examination, effective with the February 1987 administration of the CBX, and determined that "the passing score will be equivalent to the mean difficulty index before reappraisal from the last ten spring and fall bar examinations." That score was established as a scaled total score of 1440 (72 percent of the total possible scaled score).

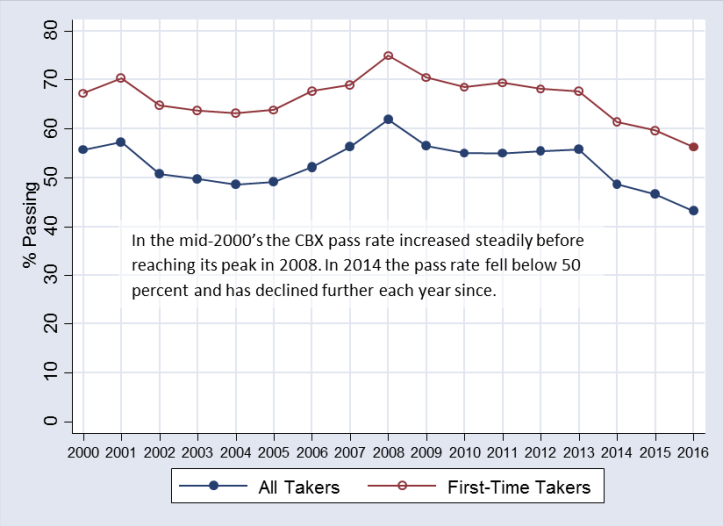
To be clear, the pass line of 1440 was not established through a standard setting study of the type described in this report. Rather, the pass line was set to maintain a relatively constant pass rate, consistent with historically observed pass rates. The minimum passing score has remained unchanged since that time.[9]

Recent Changes in the Pass Rate on the Bar Exam

California's pass line of 1440 is the second highest of all United States bar admission jurisdictions. While this pass line has remained unchanged since 1987, the pass *rate* on the CBX has risen and fallen over time: falling for a brief period in the early 2000s, then rising until it reached its peak in 2008, and declining steadily since then. Figure 1, below, shows the trend from 2000 to 2016 with the July 2016 CBX pass rate at 43 percent compared to the previous high of 62 percent in 2008.

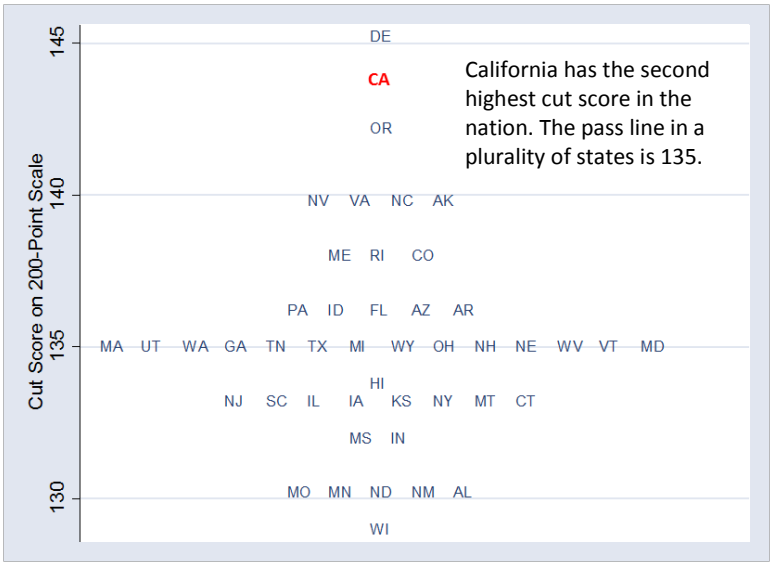
[9] In an effort to streamline the narrative, this description of the process omits various details related to committees that participated in the process, procedural issues and minor modifications to recommendations.

Figure 1. July Bar Examination Passage Rates, 2000 - 2016



Critics of the CBX point out that California's pass line is higher than that of every other state in the country except for Delaware. This high relative pass line may raise questions as to the validity of the current cut score and whether it is necessary for the purpose of protecting the public. Figure 2 shows the cut scores used by other states.

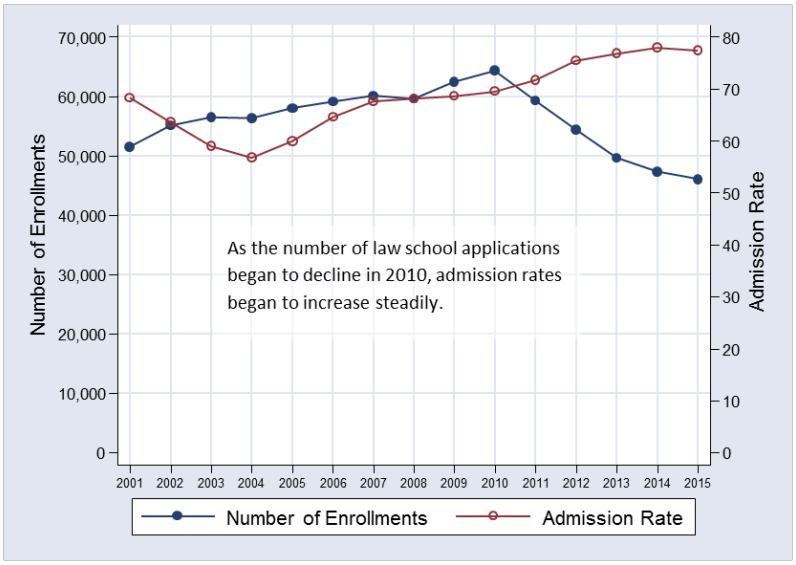
Figure 2. Passing Bar Examination Score on 200-Point Scale, by State



Because the pass rate on the CBX has risen and fallen over time without any change in the pass line, the relatively high score required for passing the exam in California cannot explain the declining pass rate over the last eight years. Moreover, many other states have experienced a similar declining trend in pass rates in recent years.

National efforts to explain these changes have focused on the precipitous drop in law school enrollment. Ms. Erica Moeser, President of the National Conference of Bar Examiners, asserts that the decline in the job prospects of newly licensed attorneys led to a decline in the number of law school applicants which, in turn, led schools to admit students with weaker academic credentials. As Figure 3 illustrates, changes in law school enrollment and corresponding admissions rates provides some support for this hypothesis.

Figure 3. National Changes in ABA Law School Enrollments and Bar Admission Rates



A 2016 proposed ABA policy exacerbated concerns regarding the decline in bar exam pass rates. In March 2016 the ABA's Council of the Section of Legal Education and Admissions to the Bar proposed amending Standard 316 to require 75 percent of a law school's students to pass a bar examination within two years of graduation for the law school to maintain its accreditation. While the proposed standard was not approved by the ABA House of Delegates, the specter of its adoption is an important contextual factor for consideration by the Board and the Supreme Court as they evaluate options for modification of the CBX pass line.

Responding to the Declining Pass Rate

In California the combination of declining law school enrollment, the looming ABA accreditation requirement, and a steady downward trend in exam pass rates came to a head after the release of the July 2016 CBX results. In the winter of 2016, the CBE determined that a thorough analysis of the issue was needed. At its December 2016 meeting the CBE took a preliminary step toward this goal by authorizing a study of the causes of the declining pass rate. The CBE contracted with Roger Bolus, Ph.D., an independent psychometrician who has been responsible for assessing the validity of each administration of the CBX for the last four years, to evaluate factors that affect the pass rate using existing Admissions data.

Adding to the momentum for evaluating these issues, in early February 2017, California law school deans wrote to the California Supreme Court urging the Court to lower the cut score on the CBX.[10] Later in February, the California Assembly Judiciary Committee held a hearing on declining exam pass rates and shortly thereafter sent a letter to the California Supreme Court urging the Court to adopt a lower pass line.[11]

In its letter, dated February 28 and quoted in the Executive Summary of this report, the Chief Justice of the California Supreme Court directed the State Bar to report back to the Court "once the investigation and all studies are concluded."[12] The report "must include a detailed summary of the investigation and findings, as well as recommendation for changes, if any, to the bar exam and/or grading, and a timeline for implementation. The State Bar's report and recommendations should be submitted to the court as soon as practicable and in no event later than December 1, 2017."

In March 2017, Dr. Bolus presented the findings of his study at a meeting of the CBE. The study looked at the relationship between pass rates and the attributes of students at different law schools. Using aggregate data, the study found that, all other things being held equal, roughly 20% of the change in July CBx scores and 17% of the change in bar passage rates could be attributed to the change in the mix of test takers between 2008 and 2016. The lack of individual, student-level data, however, limited the ability of the study to provide a full explanation of the various factors that might have contributed to thedecline in the pass rate during that period.

Subsequently, the Committee authorized and the Board of Trustees approved during its March 2017 meeting the completion of three additional, interrelated, studies regarding the CBX:

- *A Standard Setting Study*. Standard setting studies involve a formal process in which a panel of subject matter experts (SMEs), facilitated by a psychometrician in a workshop setting, evaluate the performance of exam takers and determine a pass line that is aligned with the expected level of knowledge, abilities, and skills of those exam takers.

- *A Content Validation Study*. Also relying on a panel of SMEs, content validity studies assess the alignment of examination content, in terms of breadth and depth, in relation to the expected level of knowledge, abilities, and skills of exam takers.

- *A Study of Student Performance***.**As a supplement to Dr. Bolus' study, this study aims to collect detailed, student-level data to more thoroughly evaluate any possible correlations that might exist between changes in student credentials and changes in CBX pass rates.

[10] Attached as Appendix E.
[11] Attached as Appendix D.
[12] Attached as Appendix A.

To manage these three studies and ensure appropriate oversight and communication with key stakeholders, the Bar formed a Working Group on Bar Exam Studies (Working Group) comprised of one representative from the California Supreme Court, two representatives from the Board of Trustees and two from the CBE. The Working Group is tasked with providing guidance and making administrative decisions related to the studies.

A different Working Group, composed of five deans, State Bar staff, and Dr. Bolus, was also formed to provide guidance on a study of student performance on the exam. In addition, regular updates regarding the progress of all studies underway that have been prepared for the Supreme Court are circulated to interested parties, including law school deans, the CBE, the Board of Trustees, and others. State Bar staff has also held bi-weekly conference calls since March, open to all California law school deans, to discuss progress on all three studies.

Although the three studies were launched concurrently, the Standard Setting Study was accelerated, such that should any change in the cut score be recommended by the Board of Trustees and approved by the Supreme Court as a result of that study, modification could be made in time to apply to the scoring of the July 2017 CBX.

The Standard Setting Study and a staff recommendation to circulate the study for comment was presented at a joint meeting of the A&E Committee and the CBE scheduled exclusively for this purpose on July 31, 2017. Both committees received the report and staff recommendations to consider two options – retaining the current cut score of 1440 or lowering the cut score one standard error, to 1414.,Atthe conclusion of the meeting, the A&E Committee approved releasing the report and the two options for public comment.

Following the period of public comment, staff added an additional option – a cut score of 1390 – and presented that option along with a summary of public comments and survey results to the Law School Council and a joint meeting of the A&E Committee and the CBE. On August 30, 2017, the Law School Council voted to recommend that the Supreme Court consider adoption of an interim CBX cut score within the range of 1350 to 1390, and that if a lower interim cut score is approved by the Supreme Court, the score remain in place for no less than three years.

The following day, August 31, 2017, after consideration of the staff report, comments received, and the recommendation from the Law School Council, the CBE voted to recommend maintaining the pass line at 1440 until two pending reports on the CBX – the Content Validation Study and the study of student performance – are completed and can also be considered in connection with making a recommendation relative to what the cut score should be.

Standard Setting Study and Recommendation of New Cut Score

The primary research activity for the Standard Setting Study was a two and one-half day workshop. The workshop involved a panel of 20 practicing attorneys selected to provide subject

matter expertise from a diverse body in terms of both demographics and practice types. During the workshop the panelists collectively evaluated 2,400 written essays and performance tests from the July 2016 CBX.

In preparation for the Standard Setting Study State Bar staff discussed options for a definition of minimum competence with law school deans at the April 6, 2017, meeting of the Law School Assembly. The initial resulting draft was further refined pursuant to additional feedback received from the Working Group and law school deans. During orientation and training on the first day of the workshop, the panelists discussed the concept of minimum competence that would serve as the basis for evaluating the performance levels of the essays. Standard setting panelists further modified the definition and established a baseline definition of minimum competence – also known as a Performance Level Descriptor (PLD).

It is important to note that this initial work was necessary because the State Bar did not already have a PLD that could be used for the standard setting exercise: a PLD was not used when 1440 was established as the pass line in 1987. In an effort to collect information from other jurisdictions that might inform this work, it was determined that California is not unique in lacking a PLD. No other state was found with a workable PLD that could be used.

Through the collective effort described above, the PLD was established as follows:

> "A minimally competent applicant will be able to demonstrate the following at a level that shows meaningful knowledge, skill and legal reasoning ability, but will likely provide incomplete responses that contain some errors of both fact and judgment:
>
> 1. Rudimentary knowledge of a range of legal rules and principles in a number of fields in which many practitioners come into contact. May need assistance to identify all elements or dimensions of these rules.
>
> 2. Ability to distinguish relevant from irrelevant information when assessing a particular situation in light of a given legal rule, and identify what additional information would be helpful in making the assessment.
>
> 3. Ability to explain the application of a legal rule or rules to a particular set of facts. An applicant may be minimally competent even if s/he may over or under-explain these applications, or miss some dimensions of the relationship between fact and law.
>
> 4. Formulate and communicate basic legal conclusions and recommendations in light of the law and available facts."

For each question, in the first round of activity the 20 panelists used the guidelines derived from the discussion of the PLD to classify 600 papers that they read – 30 per panelist – into three performance levels: below competence, meeting competence, and exceeding

competence.[13] Following the initial classification, in a second round of activity, the panelists further refined the performance evaluation by selecting 80 papers – four per panelist – from the first two groups (below and above minimum competence). The subset of 80 papers for each question represented the combined total of each panelist's selection of the two best of the non-competent papers and the two worst of the competent papers. These "borderline" cases provided the basis for calculating the pass line that meets the minimum competence definition.

The analysis of borderline papers to establish a new recommended cut score involved three basic steps:

1. *Mean and median cut score calculations*. Mean (average) and median (middle) scores of the papers were calculated. After assessing the difference between mean and median values, it was determined that the median value would more accurately represent the "central tendency" or average typical performance of the papers.

2. *Convergence or confidence level of cut score calculations*. Standard errors of the mean and median values were calculated to assess the degree of convergence of individual panelist's classification results and cut scores. This also allows for the representation of the statistical confidence level of the estimated mean and median values derived from the study sample.

3. *Deriving a cut score for a total scaled score*. From the initial cut scores calculated from written exam questions that were evaluated at the workshop, cut scores on the total, combined scale were derived through a process called "equipercentile linking." Equipercentile linking involves comparing two sets of scores on a common scale of cumulative percentile (from 0 to 100) that reflects the distribution of written and total scaled scores. A specific score point for written questions, after being located on the common cumulative percentile curve, is translated to a total scaled score located on the same percentile distribution curve.

Based on the above procedure the standard setting workshop arrived at a median cut score of 1439 – effectively the same as the current pass line of 1440. Following a standard statistical procedure in evaluating the degree of convergence of the cut scores, "standard errors" of mean and median were calculated to estimate a range of cut scores to provide a measure of statistical confidence. The "error" refers to the difference between the mean or median value calculated from the sample of papers evaluated in the workshop as compared to the mean or median value of the population of all exam papers. Given that a census of all 2016 Bar exam papers was neither conducted nor feasible, the standard error represents the degree of confidence to which the observed scores could be compared or generalized to the entire

[13] With each panelist reading 30 responses for each of four questions, a total of 2,400 papers were evaluated (20 X 30 X 4). The essays that panelists read were purposefully selected to present each panelist with papers that covered the full range of scores from the July, 2016 CBX. The scores, however, were not revealed to the panelists, and the papers were presented to the panelists in random order.

population of 2016 papers. As a function of sample size and the convergence of the panelists cut score values, a larger sample size and greater convergence would lead to a smaller standard error.

Because standard errors are calculated on both sides of the mean or median – in other words, the "true" value may be above or below the estimated value – the application of the standard error could result in a higher or a lower pass line. On the increase side, a cut score of 1477 and 1504 could be used, representing one and two standard errors, respectively, above the median of 1439. Conversely, on the decrease side, a cut score of 1414 and 1388, one and two standard errors, respectively, below the median could also be used to provide a full range of the possible cut scores.

This range within two standard errors above and below the average value provides a 95 percent confidence interval. This means that while 1439 is the best estimate of the "true" cut score as derived from the panelists' evaluation of the papers, one could select a value anywhere within the range from 1388 to 1504 and still have 95-percent confidence that the "true" cut score is within this range. Table 1 below shows the full range of the cut score calculations provided by Dr. Buckendahl in his technical report.

Table 1. Summary Results of Standard Setting Study

(Pass rate simulates the impact of different cut scores *if they had been used* for the 2016 CBX)

	Written Score - Mean	Combined Score – Mean (pass rate)	Written Score – Median	Combined Score – Median (pass rate)
-2 $SE_{Mean/Median}$	419	1414 (53%)	414	1388 (60%)
-1 $SE_{Mean/Median}$	424	1436 (47%)	419	1414 (53%)
Median score ($SE_{Mean/Median}$)	428 (4.47)	1451 (43%)	425 (5.60)	1439 (45%)
+1 $SE_{Mean/Median}$	432	1480 (36%)	431	1477 (37%)
+2 $SE_{Mean/Median}$	437	1504 (31%)	436	1504 (31%)

Determining where within the range the cut score should fall implicates several key policy considerations. One of these relates to the concept of two types of classification errors in statistical inference: false positive and false negative errors. Selecting a cut score above the median (or setting a higher threshold in general) is likely to lead to more false negative errors, where an applicant fails when in fact the applicant's "true" competence meets the minimum competence requirement. On the other hand, a false positive error (passing an applicant when in fact the applicant's competence does not meet the requirement) is more likely to occur when a cut score is selected below the median (or setting a lower threshold in general).

Another related policy consideration is the cost/benefit analysis of either type of error in relation to the potential tension between public protection and access to legal services. When

the threshold for entry into the practice is established at a level that is too stringent, access to justice may be negatively impacted. Conversely, a lax standard is likely to increase the risk of harm to the public.

Current bar examination grading practices demonstrate how the cost/benefit calculation has played out in California. The "second read" component of the CBX grading procedure, where examination papers from the first read that fall within a band below the cut score are re-read, represents an implicit policy position of having greater tolerance for false positive errors. The "second read" process is designed to correct false negatives by re-evaluating borderline papers that may have been, incorrectly, assessed below the cut line. Those papers that fall within a band slightly above the cut score are not passed on for second read reflecting a policy decision that is more amenable to false positive errors.[14]

In addition to error type and cost/benefit analysis, the fact that California has the second highest cut score in the nation is an important factor for the Board to consider. There is no empirical evidence available that indicates California lawyers are more competent than those in other states. Nor is there any data that suggests that a higher cut score reduces attorney misconduct. The data available to evaluate this relationship is discussed later in this report and shown in Figure 7 on page 36.

To provide the Board with additional information on which to base any recommendation that it might make to the Supreme Court, this report now summarizes the public comments received regarding the Standard Setting Study. Following that, the report will provide an impact analysis and discuss additional policy implications related to the cut score.

III. Survey Responses and Public Comments

Surveys of Attorneys and Applicants on the Cut Score

An online survey was distributed to all licensed California attorneys, including both active and inactive members, to solicit their views about the proposed options for the cut score. The survey was open for nine days from August 10 through 18 during which time nearly thirty-five thousand (34,295) attorneys responded (a response rate of about 15 percent). A slightly modified version of the survey was also distributed to bar applicants who took the July 2017 CBX. The applicant survey was open for eight days from August 11 through 18. Out of a total of more than nine thousand applicants (9,175), more than four thousand (4,188) responded (a significantly higher response rate than that of attorneys, at 46 percent).

[14] It is interesting to note that the lower bound for second read is 1390. The statistical tests applied to the Standard Setting Study data found that this lower band – 1390 – is right around two standard errors below the median, 1388.

Key questions for both surveys included options on the bar exam cut score resulting from the July 31 meeting of CBE and the A&E Committee:

- Keep the current cut score of 1440;
- Lower the cut score to 1414;
- Lower the cut score further below the recommended option of 1414; or
- Other.

Respondents were also asked to rate the importance, on a 10-point scale, of various factors that might be relevant in consideration of setting an appropriate cut score. These factors included:

- Increasing diversity of attorneys from different backgrounds;
- Increasing access to legal services for underserved populations;
- The fact that the cut score in California is the second highest in the nation;
- Maintaining the integrity of the profession;
- Protecting the public from potentially unqualified attorneys;
- Declining bar exam pass rates in California; and
- The burden of student loan debt for law school graduates unable to find gainful employment after failing the bar exam.

While not an exhaustive listing of policy issues relevant to the bar exam cut score, closed-ended responses allow for statistical reliability in assessing the importance that respondents assign to different factors. The collection of additional information related to respondents' backgrounds also provides context to better understand their views and the source of their concerns.

As shown in Table 2, the views of licensed attorneys and applicants on the cut score are sharply divergent. About 80 percent of attorney respondents are opposed to lowering the cut score, whereas only 2 percent of bar applicants hold the same view. For those attorneys who favor lowering the cut score, approximately 12 percent support the option of lowering it to 1414, and four percent lowering it further, with the remaining five percent in favor of a variety of other options, including raising the cut score, implementing the Uniform Bar Exam, eliminating the bar exam entirely, undecided or no opinion.

More than 90 percent of applicant respondents support lowering the cut score and over half (57 percent of all applicant respondents) favor lowering it below 1414.

Table 2. Survey of Attorneys and Bar Applicants on the Bar Exam Cut Score and the Relative Importance of Various Factors Related to the Cut Score

		Average Importance Rating on 10-Point Scale (1 = Not at all important - 10 = Very important)						
Option	% of Total Response	Increasing Diversity	Increasing Access	High California Cut Score	Integrity of Profession	Protecting Public Interest	Declining Pass Rate	Student Loan
Attorney (N = 34,295)								
Keep the current cut score	79.8	4.9	5.8	3.5	9.6	9.6	2.4	3.8
Lower it to 1414	11.6	7.5	8.0	6.8	8.2	8.1	6.4	7.0
Lower it further	4.2	7.7	8.3	7.7	7.1	6.9	7.7	7.9
Other	4.4	5.3	6.2	3.8	8.9	9.0	3.3	4.6
Total	100.0	5.3	6.2	4.1	9.3	9.3	3.1	4.4
Applicant (N = 4,188)								
Keep the current cut score	2.3	6.4	7.4	4.1	9.6	9.5	4.1	5.0
Lower it to 1414	36.4	7.9	8.6	7.9	8.1	7.8	8.2	8.5
Lower it further	56.7	8.4	8.9	8.1	7.6	7.2	8.8	9.0
Other	4.6	7.5	8.4	7.3	7.8	7.7	8.5	8.6
Total	100.0	8.1	8.8	7.9	7.9	7.5	8.4	8.7

Importance ratings vary significantly between attorney and applicant respondents who favor lowering the cut score. As shown in Table 2, public protection and integrity of the profession were the top factors for *attorneys* regardless of their position on the cut score. For *applicant* respondents, those who favor maintaining the current cut score rate public protection and integrity of the profession the highest also. But those applicants who support lowering the cut score rated these two factors the lowest. Instead, applicants who support lowering the cut score rated student loan debt and access as the most important factors, as well as the recent decline of bar passage rates.

Additional findings from the surveys include the following:

Attorney Employment Type**:** Attorney respondents employed in government, academia and nonprofit organizations are more likely to support lowering the current cut score. Other things being equal, those in solo practice are least likely to support lowering the cut score.

Gender**:** Female respondents, both attorneys and applicants, are more likely to express support for lowering the cut score, with a larger gender difference among attorney versus applicant respondents.

Race/Ethnicity**:** Among applicants, African Americans are more likely to support lowering the cut score (at a marginally statistically significant level) than other applicants. Other than that, there is no clear racial/ethnic distinction in applicant responses. When attorneys of color are compared to white attorneys as a whole, attorneys of color are *more* likely to favor the status quo. Broken into subgroups, the data shows a stark contrast between Asian and African American attorneys, with the former significantly more likely to support status quo whereas the

latter support lowering the cut score; there is no discernable difference between Latino and white attorney respondents.

School Type and Bar Exam Repeaters**:** Attorneys who graduated from non-ABA schools or took the bar exam more than once before passing were more likely to support lowering the cut score. Responses from applicants show no difference with respect to school types. Applicants who took the exam once or twice revealed no difference either; repeat applicants with three or more attempts, however, are more likely to be in favor of lowering the cut score. Repeat applicants as a whole are also more likely to show preference for lowering the cut score further below 1414.

In summary, the attorney and applicant surveys revealed a sharp divergence between attorneys and applicants in their assessment of the desirability of lowering the cut score. In addition, the profile of respondents shows that those who favor lowering the cut score are more likely to be women; African Americans; attorneys employed in government, academia, or nonprofit organizations; and those who had taken the bar exam more than once. Additional analysis also shows geographic variations, with both attorney and applicant respondents from Bay Area counties showing a significantly higher propensity in favor of lowering the cut score.

Public Comments and Hearings

After the joint meeting of the CBE and the A&E Committee, in which the Standard Setting Study results were discussed and staff recommendations were approved, the Standard Setting Study and corresponding options for consideration were released for public comment on August 1. The public was able to submit comments through an online public comment form or via dedicated email address. At the close of the public comment period on August 25, a total of 5,248 comments were submitted by online comment form, an additional 200 by email, and a few dozen by US mail. During two days of public hearings on August 14 in Los Angeles and August 15 in San Francisco, 25 individuals spoke on the topic.

The comments came from a variety of sources, including the general public, attorneys, bar applicants, law professors, judges, researchers and institutions, from both California and across the nation. Representatives of California law schools – ABA, California accredited, and non-accredited – also submitted letters expressing their views on the matter. For those who submitted comments via the online comment form, about 35 percent selected one of the options presented *without* providing additional narrative comments.

The comments touched on a variety of issues related to the Standard Setting Study, the bar exam, and the legal profession in general, and varied from brief statements of a few words to lengthy papers involving detailed research and analysis.

It is important to note that the comments are not a representative sample and cannot be generalized to any specific population. The online public comment form had no mechanism to prevent duplicate submissions. The number and source of comments submitted was also affected by the attorney and applicant surveys that were distributed on August 10 and 11.

Before August 10, online comments were submitted at an average rate of 100 per day and the proportion of responses in support of lowering the cut score outweighed those in favor of the status quo by three to one. During the first two days after the attorney and applicant surveys were distributed, more than 1,800 online comments were submitted, which also shifted the aggregate views with respect to cut score options. Table 3 shows that by the end of the public comment period 53 percent of online comments were in favor of the status quo, compared to 40 percent opposed.

Table 3. Selection of Cut Score Options from Online Comment Box

Option	Freq.	Percent
Keep the same*	2,789	53.1
Lower cut score**	2,073	39.5
Other***	386	7.4
Total	5,248	100.0

* Option 1 - keep current cut score of 1440 or if the option is modified.

** Option 2 - lower cut score to 1414 or if the option is modified.

*** Agreeing or disagreeing with both options.

In Appendix F the public comments are organized based on the method by which they were submitted. Online comments are grouped by the three response types noted above; individual comments within each category are then arranged in chronological order. Email submissions are the second type, followed by all others.[15]

Major Themes from Public Comments

Due to the fact that the public comment period closed on August 25, the summary of major themes provided here is mostly preliminary; lengthy comments with carefully crafted arguments have however been reviewed in detail.

To a considerable degree the seven factors presented in the attorney and applicant surveys appear frequently in various forms as major themes in the comments. For purposes of organization, key issues discussed in the comments can be categorized as follows:

- The cut score and its relationship to public protection and integrity of the profession;

[15] Due to its size – over 1,500 pages – this document is posted apart from this report in three different links on the Bar's web site.

- The bar exam and what it is designed to measure with respect to attorney competence and skills;
- Economic issues related to the supply and demand for attorney services, touching on student loan debt, with implications for access and diversity as well;
- Critique of the Standard Setting Study methodology, submitted primarily by law schools and academics in the form of letters, research papers, or public hearing presentations; and
- Reference to other states' cut scores, considering California as an unjustified outlier and citing benefits associated with lower cut scores.

Public Protection and Integrity of the Profession

While each of the broad categories outlined above reflects a wide diversity of views and individual life experiences, arguments in support of a particular position on the cut score often reflect the same concepts. Members of the public who referred to public protection and the integrity of the profession as important functions of the bar exam generally took the position that the cut score should not be lowered (with not a small number expressing support for raising it). These opinions are often expressed with an unmistakable pride that the commentators associate with the legal profession, as exemplified by the quotes below.

> "*…that California's cut score is the second highest in the nation is something to be proud of, not something to be concerned about. I have interacted with attorneys from many other states using lower cut scores. I may be biased, but I believe that in general, California attorneys are better educated and more competent than attorneys elsewhere."*
>
> *"I started community college late and was placed in remedial English and algebra....I scored a 148 on the LSAT. If I listened to arguments about how testing formats and cut scores would prevent me from achieving my dreams, I would have quit before I even tried because my credentials were not those of the "people who become lawyers." I'm glad I didn't listen and believed in myself. I worked hard to be where I am, and that is a source of pride."*

Those who take this view are prone to attribute the declining bar pass rate to the declining quality of law school students or to law schools passing unqualified students. In contrast, those who support lowering the cut score emphasize the need to consider other factors and question the relationship between the current cut score and the mission of protecting the public.

"The State Bar needs more cut score studies in order to determine with reliable empirical evidence what the proper cut score is to balance consumer safety, access to justice, and diversity in the bar."

"This high bar score does not at all reflect well on the integrity of the profession, rather it goes against the progressive stance and nature of what it means to be CALIFORNIAN. We are a state that prides itself in its forward thinking views, diversity, culture, and fairness."

Bar Exam and Measurement of Attorney Competence

"In my 15 years of practicing law, I have observed too many attorneys who lack minimal competence. It is unclear to me how lowering the cut score to allow people to practice law, who otherwise have not been able to pass, will HELP society."

"…the bar exam does a poor job determining who is actually fit to practice law. If the exam were more relevant and better tailored to determine minimal competence, then the set cut score would be much more important and should not be lowered below what California thinks is a competent threshold. However, the bar exam as it is now does not test true competence to practice law."

The two comments above represent opposing views regarding the CBX as a tool for measuring minimum competence. Those who are opposed to lowering the cut score tend to share observations of what they consider to be less than minimally competent attorneys in practice, in particular the purported decline of the writing skills of younger attorneys. Recent changes to the bar exam format - shortened from three to two days, Performance Test reduced in number from two to one, and the weight of the overall written section lowered from 65 to 50 percent – are also mentioned as evidence of the bar exam standard already being lowered.

In contrast, those in support of lowering the cut score raise questions as to the validity of the bar exam in its various aspects. At the most basic level, many point to the mere fact of the declining pass rates in recent years as an indication that something must be wrong with the exam. Going further, some of the comments criticize the CBX's heavy emphasis on rote memorization of a broad range of subject matters, a modality that is considered to be more a reflection of applicant test-taking abilities than the requisite skills of competent attorneys. The grading process of essay questions is also considered by many as flawed with testimonials from individuals who failed the bar exam attributing that failure to the "arbitrary" and "subjective" grading process that left them just a few points short of achieving the cut score level.

Independent of the role played by the CBX in testing competence, comments from those opposed to lowering the cut score often expressed the view that those who failed the exam

lacked the discipline, grit, or work ethic to succeed in passing the bar and by implication, to be a competent attorney. These comments focused on personal traits or character rather than skills or knowledge acquired through learning and training. These comments – as shown in the examples quoted below – reveal deeply embedded cultural values and often suggest inter-generational attitudinal differences, as one comment started by stating "This is typical Millenial [sic] Stuff!".

> *"I passed because of sheer grit and directed effort. I graduated as a special student from a California accredited law school while being the sole supporter of a family of 5."*

> *"To me, the declining pass rate seems to be more of a millennial problem (I say this as a millennial). Students don't want to work hard for anything, they expect it to be given to them. If they're not willing to put in the work, they should not be admitted as lawyers."*

In addition to illustrating a generational shift in values, at a personal, psychological level the interpretation of competence as a reflection of personal traits also suggests a common cognitive bias at work on both sides of the debate. "Attribution error" refers to the tendency to explain positive outcomes that we experience to our character and ability (I tried hard and passed the bar) and negative outcomes to external, uncontrollable events (I didn't pass because the exam was problematic).

Economics – Supply and Demand of Legal Services, Access and Diversity, Student Loan Debt

Each of the seemingly disparate topics summarized in this section deserves its own treatment; the synthesis presented here does not pretend to do justice to the complexity of the interconnected issues. There is, nonetheless, a common thread of economic or market-related concerns raised when issues of access, diversity, and student loan debt are discussed in connection with bar pass rates and cut score. This is also evident in attorney and applicant surveys, as discussed above, in which the ratings on access, diversity, and student debt are closely associated with one another.

Many who provided comments in favor of keeping the current cut score asserted that there is no shortage of attorneys in California as a primary argument against lowering the cut score. In fact, references to the job market for attorneys appeared in nearly a quarter of all online comments taking a position in support of keeping the current cut score.

> *"To be clear we DO NOT have a shortage of lawyers in this state but we do continue to have a shortage of high-quality legal jobs."*

*"Consider: there is, most assuredly, no shortage of attorneys in the state -- judges will tell you that there is a shortage of *effective* attorneys in their courtrooms. Lowering standards will not address this problem, but merely make it worse."*

For those in favor of lowering the cut score, economic issues – whether employment, job market, or student loan debt – are often presented in narratives involving difficult choices and personal sacrifices, lost opportunity for gainful employment and service to the community, crushing financial burden from student loan debt and mental agony in dealing with all of these challenges. The emotional intensity appears to be captured in applicant survey results showing that access and student loan debt are rated at the top as most important factors for applicants in favor of lowering the cut score.

"…it has been an extreme financial burden as I have spent over $20,000 dollars in bar fees, accommodations, and resources to prepare for the bar. Moreover, my employment opportunities and earning potential are on standby until I pass the bar."

"*During my final semester of law school, I interviewed for legal jobs across the country. When time came to register for a bar exam prep course, I was faced with a quandary: do I bet on the California bar exam, in hopes that a California employer says yes to me? Or do I study for a bar exam with greater reciprocity, such as the New York State bar exam, to allow me to start practicing immediately in whatever state I may land a job? Ultimately, I chose to take the California state bar exam, which cost me several promising prospects."*

<u>Reference to Common Cut Score Ranges in Other States</u>

Many public comments point to other states' cut scores as evidence that California's pass line is unjustified and extraordinarily high. These comments often cite the benefits associated with a lower cut score, primarily by increasing diversity of the profession and access to legal services. As to the specific cut score below the current level, two stood out as preferred by the vast majority: 1390 (rounded from 1390, or 139 based on a 200-point scale comparable to other states) and 1350, the most common cut score in other states.

Preference for 1390

Two main arguments are often presented as justifying a cut score of 1390:

1. it is within the confidence level of the cut score range from the Standard Setting Study;
2. it is consistent with the score range from 1390 to 1439 for written exams that are currently eligible for second read (some comments assert that papers within this range

are often comparable in their quality, aside from grader subjectivity, calling into question the reliability of differentiating performance within this range).

It is true that from the Standard Setting Study result, 1390 lies at the lower end of the confidence interval two standard errors away from the median. It should be pointed out, however, the further a score point within the confidence interval deviates away from the center (1440), the less confidence one would attach to the score due to the bell-shaped distribution. Following the same reasoning, one would have similar confidence in preferring a cut score to the right of the center at a higher cut score level. However, while statistical evidence presents a framework to evaluate the degree of certainty in selecting a particular cut score level, the decision must ultimately be informed by considering all of the policy implications.

It is not true, though, that the exams currently eligible for a second read are indistinguishable from those that pass. Historically the proportion of applicants who passed the bar after the second read has remained fairly constant at around 15 percent of the total second-read papers. The closer a first-read score is to the pass line, the greater the likelihood that it would pass after the second read. For example, for applicants whose scores fell between 1430 and 1439 on the 2016 July exam, approximately 35 percent passed after second read. That proportion falls by almost half for each 10-point decline within the second-read zone. For those at the bottom end of the zone, between 1390 and 1400, less than two percent passed after the second read.

Preference for 1330 or 1350

Those who advocate lowering the cut score below 1390 often point to the fact that California has the second highest cut score in the nation. Even lowering it to 1390 would leave California with one of the highest cut scores in the country, particularly after Oregon and Nevada, two states at the higher end, recently decided to lower their cut scores. These comments also tend to point out that in states with lower cut scores, including New York with a cut score of 1350, there is no evidence that the competence of attorneys is compromised or that more disciplinary actions are associated with lower cut scores. This last statement is borne out in the brief analysis contained in this report and shown in Figure 7 on page 36.

As the range between 1330 and 1350 is most common in other states, a significant number of proponents for lowering the cut score in California advocate this range by appealing to notions of the “wisdom of the crowd” or “crowd sourcing” as validity evidence regarding attorneys’ minimum competence.

Methodological Critiques of the Standard Setting Study

A third category of public comment stands in contrast to the survey results and the majority of the public comment narratives. This third category, primarily from law school representatives and other academics, is more technical in nature and tends to focus on specific methodological and procedural issues. Because the assertions contained within these comments call into question the very validity of the Standard Setting Study, they are addressed point-by-point below.

A number of these comments criticized the statement of minimum competence, the PLD, that was used for evaluating papers. These concerns pointed to the amount of time devoted to the development of the PLD and the amount of detail contained in the definition of minimum competence. These critiques also focused on the absence of detailed grading rubrics for each exam question presented at the workshop. Related criticisms pointed to the amount of time spent with the panelists at the workshop to discuss and refine the PLD, and a number of comments pointed to one particular member of the panel who had extensive experience grading bar exam papers and who was perceived to be excessively influential during discussions.

Many of these same issues are mentioned by one of the outside reviewers, Dr. Mary Pitoniak. In her detailed overview and critique of the process, Dr. Pitoniak noted that she would have preferred for more time to be been spent defining the PLD and also noted the apparently outsized influence of one panel member.

In addition to the various procedural issues noted above, these comments also focused on the measure of central tendency employed in the study. Two comments, in particular, suggest that the cut score results were too widely dispersed and that the distribution pattern lacks the characteristics of a typical normal distribution as the basis for calculating an overall mean or median cut score value.

1. Panel Composition and Panel Dynamic

One of the key elements of the Analytical Judgment Method (AJM) used for this study and for standard-setting studies in general, is an emphasis on the transparency and inclusiveness of the process. From the initial stage of recruiting panel members to the participation in the study, the State Bar reached out to various stakeholders to solicit nominations. The final selection of panelists made by the Supreme Court was drawn from almost 40 nominees, and consisted of practitioners from a broad representation of backgrounds in terms of practice area, experience, demographics, and geographic distribution.

Broad stakeholder participation in the panelist nomination process is an essential element of the process for ensuring procedural validity of the AJM. In this instance, such participation was sought from all three California law school types. Although Dr. Tracy Montez pointed out that the Department of Consumer Affairs attempts to avoid using educators on panels for standard setting studies to avoid potential conflicts of interest, two educators were selected to serve on the panel.

The decision to include the educators was made to balance the desire for inclusivity with the need for analytic rigor and procedural fidelity. Given the size of the 20-member panel, which typically ranges from 15 to 30 depending on the particular setting and specific study design, coupled with the large number of exam papers reviewed, we believe that the educators made important contributions to the study without skewing the final results.

As a follow up to the Standard Setting workshop, Bar staff conducted interviews with panelists to solicit their feedback about the workshop process. One of the questions specifically addressed whether they felt that they had the opportunity to fully participate in the discussions, and whether any individual members had more influence than others in steering the direction of the discussions. Of the 15 panelists interviewed, all but two members indicated their awareness that one of the members seemed more vocal due to his in-depth knowledge of bar exam grading.

As to whether their views had been influenced by this panelist or whether this panelist affected the dynamic of the group discussions in an undesirable manner, the common response was to dismiss his influence. The sentiment most frequently expressed in these interviews was that the panelists are all independent-thinking and successful attorneys, and that diverse views are helpful: when people make a forceful argument, they still have to convince others. This reaction is consistent with the literature on discourse analysis which emphasizes the role of social status differentials in a group setting.

2. Minimum Competence Definition and Grading Rubric

Two early versions of the PLD, excerpted and modified from bar exam grading guidelines, were first presented for discussion with law school deans at the Law School Council meeting in April 2017. The topic was on the agenda several times thereafter during the bi-weekly conference calls with law school deans that the Bar has held since April. The final version of the PLD reflected the combined input of law schools, Supreme Court staff, and the study psychometrician, Dr. Chad Buckendahl, whose focus was on ensuring that the language used in the statement complied with best practice in standard-setting procedures.

With the PLD serving as the criterion to evaluate the performance of exam papers, grading guidelines or rubrics are intended to provide more concrete information to panelists as a way

of focusing on specific elements or traits of the exam. Used in conjunction with the more general definition of PLD, the purpose is to ensure that the papers are evaluated on the basis of a consensus definition as *criteria-referenced* activity rather than *norm-referenced* activity that relies on the relative performance of a set of papers being evaluated.

The original research design actually included individual rubrics for the evaluation of each question. The decision to use a generalized rubric instead of question-specific rubrics was based on concerns expressed by law school deans that rubrics at the item level would bias the panelists. This modification in the method represents a preference rather than a flaw in the method and balances the input of stakeholders without threatening the validity of the research.

The critiques of the PLD, similarly, represent preferences rather than substantive flaws in the method. The PLD was directly related to the judgments that panelists were asked to make on the examination, specifically the written portion. As a policy statement, the language is intended to provide a profile of what a minimally competent candidate knows and is able to do. This purpose was accomplished through the initial draft and discussion among the panelists. The appropriate question for the panelists was whether the samples they reviewed were characteristic of the performance of minimal competency as guided by the PLD, and this was accomplished during the workshop.

3. Measurement Issues

The most detailed critiques focused on measurement came primarily from two sources: Professor Deborah Merritt of Ohio State University, Moritz College of Law, and Dr. Benjamin Nyblade, director of the Empirical Research Group at UCLA School of Law. Professor Merritt and Dr. Nyblade point to some of the same procedural issues discussed above but also assert that the statistical patterns presented in their critiques demonstrate a "fatal flaw" in the study results.

To summarize the response to their critique in advance of providing the detail that supports it: the analyses and conclusions of Professor Merritt and Dr. Nyblade are correct in terms of the dispersion of the data that they reviewed – but they were looking at data that does not fully reflect the entire process that generated the final result of the cut score range. Instead of reviewing the entire range of panelist responses – 320 data points used to derive the estimated mean and median – these two critiques looked at only the final 20 data points that were calculated from the larger sample.

To reiterate how the final cut score recommendations were derived:

- In the first step of the Standard Setting Study, the 20 panelists evaluated 30 essays for each of four questions – a total of 120 papers per panelist for a total of 2,400 essays;

- Panelists then selected four essays for each question – two marginally above competent, and two marginally below;
- Thus, the four different questions multiplied by 20 panelists and then multiplied by the four essays selected yielded 320 data points – 160 on each side of the border of minimum competence.

In Dr. Nyblade's comment, the presentation of these data points is compressed into a graph that sets the minimum and maximum values equal to the range of the 20 average values derived from the 320 data points, ranging from 380 to 460. This gives the appearance of cutoff scores that are, in Dr. Nyblade's words, "all over the map." When looking at the entire process, however, including each step of selection, and all of the data points used to calculate the 20 average values, the problem appears to be more akin to looking at the *inset* of a map and noting how broadly dispersed the landmarks are within it, without recognizing that this inset is on a different scale from the larger map under consideration.

It is important to keep in mind the multiple steps that the panelists followed in evaluating the papers, the meaning of the data derived from each step, and to interpret each in the proper context. Before the panelists started to formally evaluate the papers, they were first given ten papers as an exercise. This occurred shortly after the conclusion of the PLD discussions to ensure the application of the consensus concept of minimum competence to exam papers as an independent judgment activity.

With further discussion after this first round of exercise activity, panelists began to review 30 papers for the first question. The 30 papers were first classified into three performance levels: below, meeting, and exceeding minimum competence. Analogous to a calibration session to further refine the PLD, the panelists honed in on "borderline" cases by selecting two cases from the below-competence group and two from the meeting-competence group; the exceeding-competence papers were not included in this step.

Following the steps described above, each panel member classified 120 papers, for a total of 2,400 papers for the entire panel. The second part of the process produced rating results for 320 papers (4 papers from each question, 4 questions from each panelist, yielding 320 (4 × 4 × 20). These 320 data points provide the basis for calculating the cut score recommendations reflected in the study results.

The three graphs below show the distributional patterns of these 320 data points from different angles. Figure 4 arranges the 320 papers into two groups: best of non-competent and worst of competent papers. It shows where along the score ranges (vertical axis) the papers are clustered, range of the middle 50 percent of the papers (rectangle boxes at the center of the leaves), and the median value within each group (white dots at the center). Scores for the first

group of the best of non-competent papers are clustered around 50 points (mode of the distribution), with median value located at approximately 60 score points. For the second group to the right, the scores are clustered near the center of the leaf around 60, with median value located at approximately 65. Each group consists of scores from 160 papers.

Figure 4. Score Distribution Patterns of Borderline Papers

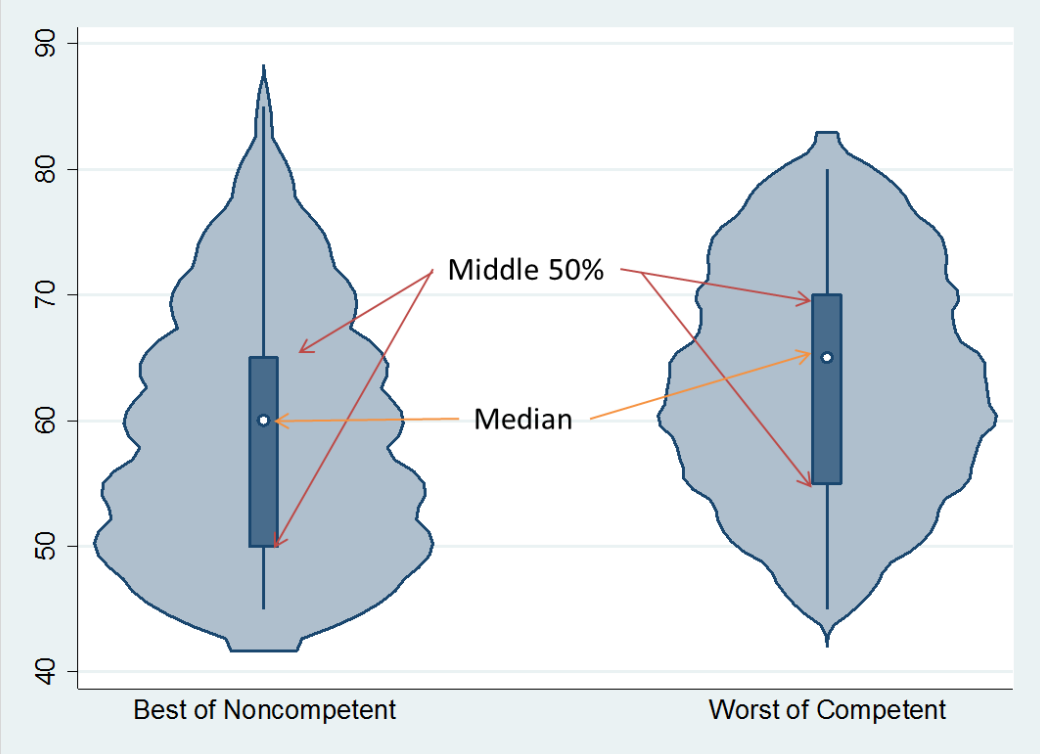
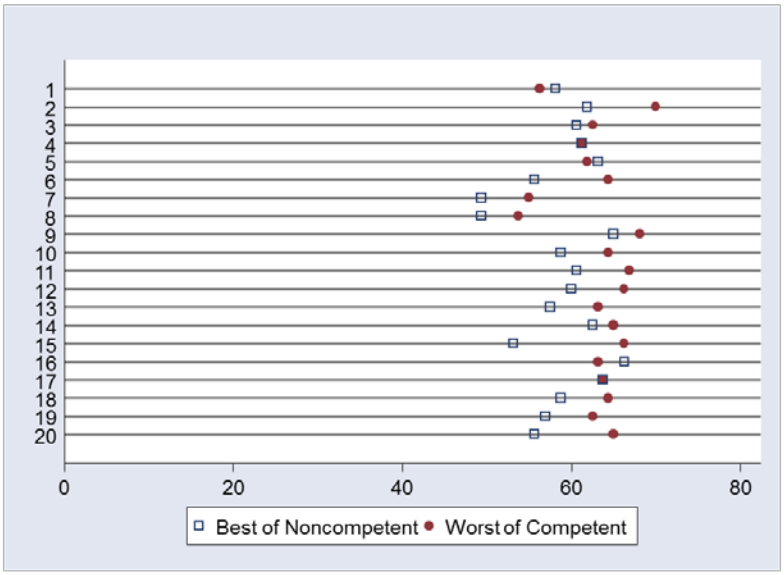


Figure 5 shows the average scores for the two groups of borderline papers for each panel member. Red dots represent the worst of competent papers and hollow squares represent the best of non-competent papers. The average scores for the worst of competent papers are expected to be higher than those for the best of non-competent papers (red dots located to the right of hollow squares). The distance between the two data points for each panelist could serve as an indication of the degree of precision in identifying borderline cases at the margin. The mid-points between the red dots and hollow squares represent the cut score at the level of panelist.

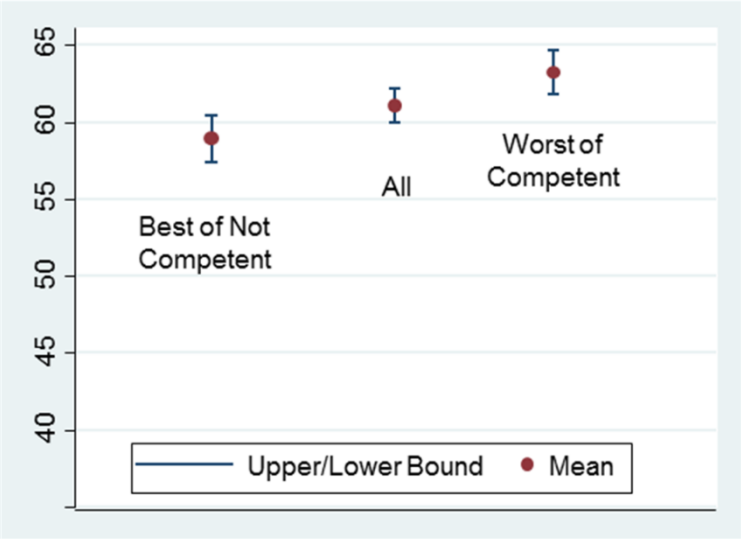
It should be noted that, with the exception of two panelists whose mid-point falls around 50, mid-points fall consistently slightly above (to the right on the horizontal axis) 60. At this level of analysis each dot represents 8 data points – the two papers judged to be the worst of the competent for each of four essay exams – and each square represents eight data points – the two papers judged to be the best of the non-competent for each of four essay exams.

Figure 5. Average Scores of Borderline Cases, by Panel Member



Aggregating the data still further, Figure 6 shows the average value for each of the two groups of borderline cases, as well as the summary of all borderline cases. The red dots represent the average value, with the length of the vertical line through the dot representing the margin of error around the mean value. The average score for all borderline papers is derived as 61 score points, with a lower bound in the confidence band at 59.98 and an upper bound at 62.15. Taking 61 as the overall average for borderline papers and aggregating to a full exam consisting of 5 essays and 1 Performance Test (given twice the weight of the essays), the total score of 427 (61 X 7) approximates the mean cut score of 428 reported in the final report of the study.

Figure 6. Overall Score Point Averages of Borderline Cases



The analysis at each level of aggregation demonstrates the convergence of the scores derived from the independent judgment of the 20 panelists, despite a small number of outlier cases. It also provides indirect confirmation of the consensus among panelists in their discussion of the PLD, and the consistent application of grading guidelines to the performance of the papers.

Instead of looking at the entire sample of 320 data points that reflect the result of a two-step, calibrated evaluation process, critics of the study chose to focus on 20 data points to assert the absence of a valid distribution of the data, thus calling into question the validity of the entire study. Admittedly the multiple steps in the standard-setting process might have caused confusion regarding the proper unit of analysis. Regardless of the reason for the confusion, this explanation should suffice to minimize any misunderstanding regarding these concerns about the methodology.

In her summary of the entire standard setting process, Dr. Pitoniak concludes that:

> "in my opinion there were no fatal flaws. The panel-recommended passing score, and the recommendations for adjusting it made by Dr. Buckendahl, represent credible information for the Supreme Court to consider when they make their policy decision."

Dr. Montez's summary of the standard setting process was less detailed and focused more on compliance with procedural requirements of the methodology. Dr. Montez's frames the observation and review of the workshop as an evaluation "to determine whether the standard setting procedure meets professional guidelines and technical standards outlined in the *Standards for Educational and Psychological Testing*." Dr. Montez concluded the summary of her report noting that:

> "The State Bar and ACS [Dr. Chad Buckendahl's consulting firm], appear to adhere to professional guidelines and technical standards, but also recognize that additional strategies can be implemented to further add evidence supporting the pass/fail decision based on CBE [CBX] performance."

IV. Policy Issues Related to the Pass Line

Though the CBX is an essential component of admission to the practice of law in California, no exam can perfectly predict the performance of applicants, safeguard the public with certainty, or ensure an adequate supply of attorneys to meet the demand for legal services. While the admissions process addresses some of these factors – through the moral character evaluation and Multistate Professional Responsibility Examination, for example, others must be addressed as matters of public policy.

The following section of the report assesses the potential impact of changing the CBX cut score by looking at the differential impact of the cut score on various applicant groups. This analysis looks specifically at the simulated impact of cut score of 1414 and 1390, reflecting the full range of available pass line options that fall within two standard errors of the Standard Setting Study median.[16]

The Simulated Impact of Changing the Cut Score in California

The implications of selecting a modified pass line of 1414 or 1390 are shown in Table 4 and summarized below. This discussion describes what the impact *would have been* if the pass line had been altered and applied to the 2008 and 2016 administrations of the CBX. Because a decrease in the cut score will necessarily increase the number of applicants who passed the exam on the two samples to which we apply this simulation, the issues examined here relate to the *magnitude* of the changes at different scores and the *differential impact* of the changes on different sub-populations of the total applicant pool.

- Simulating the impact of a cut score of 1414, the total percentage of applicants who passed the exam would have increased from 61.9 percent to 65.5 percent in 2008 and would have increased from 43.3 percent to 46.8 percent in 2016;
- If the cut score had been 1390 for those two administrations of the exam, the total percentage of applicants who passed the exam would have been 69.9 percent in 2008 and 52.1 percent in 2016;
- Simulating the different pass points for men and women taking the exam:
    - A 1414 pass line would have resulted in an increase from 60.9 percent to 64.3 for men taking the exam in 2008 and an increase from 44 percent to 47.4 percent of men taking the exam in 2016;
    - Applied to women, a 1414 pass line would have increased the percentage passing from 63 percent in 2008 to 66.9 percent and would have increased the percentage passing from 42.5 percent to 46.1 percent in 2016;
    - A pass line of 1390 would have increased the percentage of men passing the exam to 68.9 percent in 2008 and to 53.1 percent in 2016;
    - For women, a 1390 pass line would have increased the percentage passing the exam to 71 percent in 2008 and to 51.2 percent in 2016.

Before presenting simulation data outlining the impact that different pass points would have by various racial and ethnic categories it is important to make a brief comment on the math. Groups with *lower* pass rates at the current pass line will almost necessarily show

[16] Although the discussion here looks only at the simulated impacts of cut scores of 1414 and 1390, in response to a public records request, Appendix G has been added to this report showing the simulated impact of cut scores of 1330, 1350, 1390 and 1414.

greater percentage increases in pass rates as the pass line is lowered than groups with *higher* pass rates at the current pass line. As a result, African Americans and Latinos show the largest percentage increase in pass rates despite larger raw numbers of whites and Asians who would pass at a lower pass line.

- Thus at a pass line of 1414:
    - 38.1 percent of African Americans taking the exam in 2008 and 23.1 percent taking the exam in 2016 would have passed, increases of 10.4 and 12.5 percent respectively;
    - 53.2 percent of Latinos taking the exam in 2008 and 37.5 percent taking the exam in 2016 would have passed, increases of 8.8 and 10.6 percent respectively;
    - 59.9 percent of Asians taking the exam in 2008 and 40.5 percent taking the exam in 2016 would have passed, increases of 6.4 and 8.7 percent respectively;
    - 71. 6 percent of whites taking the exam in 2008 and 54.9 percent taking the exam in 2016 would have passed, increases of 5.2 and 7.2 percent respectively.
- In addition to looking at the simulated changes within groups at different pass lines, we can also look at the changes in the relationships among groups, or the disparities in pass rates among different groups. Looking at the 2008 administration of the CBX:
    - At the current pass line of 1440, the pass rate of whites is 97.1 percent higher than that of African Americans, a gap of 33.5 percentage points. At a simulated pass line of 1414, the gap in the pass rates between whites and African Americans declines to 87.9 percent, and at 1390, the gap in the pass rate between whites and African Americans declines to 65.6 percent, a difference of 27.9 percentage points;
    - Looking at the 2016 administration of the CBX, the gaps are much larger between whites and African Americans but the diminution of the disparity is also significant at the simulated pass lines. At the 1440 pass line, whites passed the bar exam at a rate one-and-a-half times greater than that of African Americans (149 percent). Simulating the impact of lower cut scores, that gap would have diminished to 138 percent at 1414 and to 108 percent at 1390.

Table 4. Simulated Impact of Pass Rates at Different Cut Scores

		2008 CBX			2016 CBX		
Cut Score		**1390**	**1414**	**1440**	**1390**	**1414**	**1440**
Total	# Passing	6,017	5,642	5,329	4,010	3,598	3,332
	% Passing	69.9%	65.5%	61.9%	52.1%	46.8%	43.3%
	% Increase*	*12.9%*	*5.9%*		*20.3%*	*8.0%*	
First Time	# Passing	5,078	4,870	4,682	3,317	3,066	2,896
	% Passing	81.4%	78.0%	75.0%	64.5%	59.6%	56.3%
	*% Increase**	*8.5%*	*4.0%*		*14.5%*	*5.9%*	
Repeat	# Passing	939	772	647	693	532	436
	% Passing	39.7%	32.6%	27.3%	27.2%	20.9%	17.1%
	*% Increase**	*45.1%*	*19.3%*		*58.9%*	*22.0%*	
Male	# Passing	3,121	2,911	2,756	1,970	1,760	1,635
	% Passing	68.9%	64.3%	60.9%	53.1%	47.4%	44.0%
	*% Increase**	*13.2%*	*5.6%*		*20.5%*	*7.6%*	
Female	# Passing	2,890	2,726	2,568	2,005	1,805	1,665
	% Passing	71.0%	66.9%	63.0%	51.2%	46.1%	42.5%
	*% Increase**	*12.5%*	*6.2%*		*20.4%*	*8.4%*	
Asian	# Passing	1,205	1,113	1,046	835	735	676
	% Passing	64.8%	59.9%	56.3%	46.1%	40.5%	37.3%
	*% Increase**	*15.2%*	*6.4%*		*23.5%*	*8.7%*	
Black	# Passing	215	181	164	146	117	104
	% Passing	45.3%	38.1%	34.5%	28.9%	23.1%	20.6%
	*% Increase**	*31.1%*	*10.4%*		*40.4%*	*12.5%*	
Hispanic	# Passing	471	432	397	478	419	379
	% Passing	58.0%	53.2%	48.9%	42.8%	37.5%	33.9%
	*% Increase**	*18.6%*	*8.8%*		*26.1%*	*10.6%*	
White	# Passing	3,765	3,570	3,392	2,369	2,165	2,019
	% Passing	75.5%	71.6%	68.0%	60.1%	54.9%	51.2%
	*% Increase**	*11.0%*	*5.2%*		*17.3%*	*7.2%*	
Other	# Passing	71	67	60	66	56	52
	% Passing	57.3%	54.0%	48.4%	44.6%	37.8%	35.1%
	*% Increase**	*18.3%*	*11.7%*		*26.9%*	*7.7%*	
ABA	# Passing	3,767	3,571	3,415	2,629	2,387	2,231
	% Passing	82.3%	78.0%	74.6%	63.8%	57.9%	54.2%
	*% Increase**	*10.3%*	*4.6%*		*17.8%*	*7.0%*	
CA Accredited	# Passing	265	225	196	169	131	100
	% Passing	35.6%	30.2%	26.3%	21.9%	17.0%	13.0%
	*% Increase**	*35.2%*	*14.8%*		*69.0%*	*31.0%*	
Registered	# Passing	107	88	76	44	38	35
	% Passing	33.5%	27.6%	23.8%	16.2%	14.0%	12.9%
	*% Increase**	*40.8%*	*15.8%*		*25.7%*	*8.6%*	
Out of State	# Passing	1,369	1,307	1,242	801	730	685
	% Passing	73.2%	69.9%	66.4%	56.5%	51.5%	48.3%
	*% Increase**	*10.2%*	*5.2%*		*16.9%*	*6.6%*	

** Percent increase of the number of applicants that would have passed under each simulated cut score level relative to the number of passing applicants under the current cut score of 1440.*

Issues of Diversity and Access

Lowering the cut score on the bar exam and, as a result, increasing the number of attorneys in California would not, by itself, increase the availability of attorney services for those who most need them. Access to legal services depends on where attorneys choose to practice, the type of law they choose to practice, the cost of legal services, and other factors beyond simply increasing the pool of attorneys who practice law in California.

That said, data on the types of law practiced by attorneys of different backgrounds suggests that attorneys of color and women tend to practice public-interest and non-profit law more often than white men. In a January 2017 survey conducted by the State Bar of California, just under four percent of white respondents indicated that they worked in the non-profit sector. In contrast, over eight percent of Latino attorneys, six percent of African American attorneys, and 6.3 percent of Asian attorneys indicated that they worked in the non-profit sector.

The survey data show similar disparities across different attorney groups in the likelihood of working in government. Women are almost twice as likely to work in government as men – 23 percent of female attorneys surveyed indicated that they worked in government compared to 13 percent of male attorneys. Looking at different racial / ethnic groups in the government sector, whites are the least likely among the groups identified to work in government – 16 percent of white attorneys surveyed, compared with 22 percent of Latino and Asian attorneys and 32 percent of African American attorneys. Though government work can encompass various forms, many of these positions are essential to improving access to justice including public defenders, juvenile justice attorneys, and self-help attorneys in the courts.

These findings align with research conducted by Wendy Espeland, Associate Professor of Sociology at Northwestern University and Michael Saunder, Associate Professor of Sociology at the University of Iowa who found that law school graduates of color are more likely to start their law careers in government and public interest law than their white counterparts.[17] As a consequence, issues of diversity appear to be linked to issues of access.

The need for attorneys outside of the private sector has been documented by the Legal Services Corporation (LSC) in a report describing the negative impacts that the justice gap has on varying communities.[18] In their work, LSC defines the justice gap as the "difference between the civil legal needs of low-income Americans and the resources available to meet those needs." Specifically, the lack of legal services in civil matters has the greatest impact on seniors, rural

[17] See "Rankings and Diversity," *Review of Law and Social Justice*, 2009, 18, 587-608.
[18] Legal Services Corporation. 2017. The Justice Gap: Measuring Unmet Civil Legal Needs of Low-income Americans. Prepared by NORC at the University of Chicago for Legal Services Corporation. Washington, DC.

residents, veterans, persons with disabilities, parents of children under 18 and survivors of domestic violence or sexual assault.

Surveys of 2,000 adults living at or below 125 percent of the Federal Poverty Level reveal that while 71 percent of these households had at least one civil legal problem in the past year, only 20 percent sought professional legal help for these legal problems. Low-income Americans are the least likely to understand the complexity of the legal system or trust it as an institution to which they can turn for assistance. When confronted with a civil legal problem, many low-income Americans may not even realize that the problem has a potential legal remedy.

Where low-income Americans do find their way to the justice system, increasingly they are unrepresented. In California an estimated 81 percent of unlawful detainer proceedings include at least one party who is self-represented; in domestic violence proceedings 90 percent of proceedings involve at least one unrepresented party.[19] Similarly, family law proceedings are increasingly managed by litigants without representation.[20]

Beyond the documented tendency of California attorneys of color to work outside of corporate law, diversity in the legal field is also important to the extent that public perception of the law instils – or undermines – confidence in the legal system. New York State Court of Appeals Judge Jenny Rivera asserts that there are three goals of diversity in law: establishing a profession that represents the broad diversity of the population it serves; providing proof that the legal system does not have barriers based on race, ethnicity, gender; increasing public confidence in the administration of justice; and to promulgate the belief that the system is fair.[21]

Diversity advocates argue that increased public trust and confidence in the legal system also contributes to increased compliance with the law. The Strategic Plan of California's Judicial Branch includes Access, Fairness, and Diversity as the first of its seven Strategic Goals[22] and diversity considerations fall squarely within the State Bar's newly adopted mission statement:

> "The State Bar of California's mission is to protect the public and includes the primary functions of licensing, regulation and discipline of attorneys; the advancement of the ethical and competent practice of law; and support of efforts for greater access to, and inclusion in, the legal system."

[19] See "The Justice Gap: A Crisis on the Courthouse Steps," http://www.calbar.ca.gov/Portals/0/documents/accessJustice/2015_JusticeGapFund_FactsandFigures.pdf

[20] Bonnie Hough, "Self-Represented Litigants in Family Law: The Response of California's Courts," *California Law Review*, February, 2010.

[21] Rivera, Jenny. (2016) Diversity and the Law. *Hofstra Law Review,* 1271-1286.

[22] Judicial Council of California website (2017). Retrieved from: http://www.courts.ca.gov/3045.htm.

Despite these worthy goals, articulated for a number of years by the Judicial Branch, California's attorney population remains disproportionately white and male.

Comparing the demographic make-up of California with the findings of the 2017 survey of licensed attorneys in California shows that while Latinos make up 35.4 percent of Californians over the age of 18, they represent less than five percent of California's licensed attorneys. African Americans make up 5.9 percent of the state's population over 18 years of age but account for less than two percent of licensed attorneys while Asians comprise 13.8 percent of the population over 18 but just under six percent of licensed attorneys.

While the root causes of disproportionate rates of passage are beyond the scope of this report, it is clear that applicants of color pass the bar exam at rates that are disproportionate to those of their white counterparts. This impact, when combined with disproportionately lower numbers of people of color in the pipeline to higher education and law school, has resulted in a pool of licensed attorneys in California that does not reflect the population of the state.

As Table 4 on page 32 shows, reductions in the CBX cut score would not eliminate disparities in the pass rate among different groups of applicants. Reductions in the cut score would, however, reduce those disparities and, if the patterns of career choice hold for new attorneys, might also improve access by licensing more attorneys with a propensity to work in the non-profit and government sectors.

Public Protection and the Pass Line

Many of the public comments and survey responses from people who expressed a preference for maintaining the current cut score of 1440 also indicated that the integrity of the profession and public protection were the most important issues in their consideration of the cut score. Where public protection is conceptualized as ensuring that licensed attorneys are minimally competent, the bar exam must serve this function and the cut score should prevent those who lack minimum competence from becoming licensed. It is less clear, however, what role the exam and cut score play in other forms of public protection. To the extent that higher cut scores reduce the number of licensed attorneys in a given jurisdiction, a high cut score may in fact undermine access to legal services – an important form of public protection.

One area where we can examine the relationship between public protection and the pass line empirically is in a simple evaluation of attorney discipline and the pass line in different states. The scatter plot in Figure 7 shows attorney discipline data from the 2015 ABA Survey on Lawyer Discipline Systems and evaluates it in relationship to the cut scores in 44 states for which data were available.

What the scatter plot shows is that attorney discipline – as measured by private and public discipline per thousand attorneys – appears to have no relationship to the cut score. With so many states using 135 for their cut score, the details of the Figure can be somewhat difficult to tease out. The big picture, however, is clear. At a cut score of 135 the rate of attorney discipline ranges from a low of 1.9 per thousand in West Virginia to 7.9 per thousand in Tennessee. Looking across the entire range of cut scores we see strikingly similar rates of attorney discipline in states with cut scores from 130 – Alabama – all the way to 145 – Delaware. California's rate of discipline (2.6 per thousand) is less than a third the rate of Deleware (4.7 per thousand).

Given the vast differences in the operation of different states' attorney discipline systems, these discipline numbers should be read with caution. But based on the data available, it appears unlikely that changing the cut score would have any impact on the incidence of attorney misconduct.

Figure 7. Relationship between Cut Score and Attorney Discipline

V. Cut Score Options

In the report submitted to the CBE at its meeting on July 31, staff presented the Standard Setting Study report and, based on the study, presented two options for consideration:

1. maintain the current cut score of 1440, or
2. reduce the cut score to 1414 on an interim basis.

After reviewing comments submitted during the public comment period, staff now presents the Board with a third option: reduce the cut score to 1390.

Staff presents this third option for three reasons. First, the option is responsive to public comments urging further reduction to 1390.

A second reason for presenting the option of 1390 is because this value falls within the confidence interval of two standard errors of the median estimated cut score determined from the Standard Setting Study. Because measurement error is inevitable when calculating values based on a sample, a confidence interval provides a statistical test of the likelihood that the correct value falls within a particular range. In the case of the Standard Setting Study, the score of 1390 falls within the margin of error, the range within which it can be said that there is a very strong likelihood – greater than 95 percent at two standard errors – that the "true" cut score is between 1388 and 1504.

Confidence intervals are especially useful in situations where there is uncertainty about the exact number and provide decision-makers with a range of options within which they can attempt to balance competing values. Thus, while in this case a cut score two standard errors in either direction would be defensible *analytically*, when balanced against other factors, a cut score of 1390 appears to be a viable option.

The third reason for offering 1390 as an option for consideration relates to the issues of access and diversity summarized above. Despite the disproportionate pass rates across all potential cut scores, 1390 would decrease the disparity in pass rates without straying beyond a statistically defensible margin. This argument applies to a reduction of the cut score to 1414 or to 1390, but the impact is the largest at 1390.

In response to these same recommendations, the Law School Council voted to recommend a cut score within a range of 1350 to 1390. The CBE voted to recommend that the cut score remain at 1440 pending the completion of additional research.